

DISCRIMINATION BETWEEN STATIONARY GUASSIAN PROCESSES, LARGE SAMPLE RESULTS

BY

WILL GERSCH

TECHNICAL REPORT NO. 30

JANUARY 1977

PREPARED UNDER CONTRACT N00014-75-C-0442

(NR-042-034)

OFFICE OF NAVAL RESEARCH

THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA



DISCRIMINATION BETWEEN STATIONARY GUASSIAN PROCESSES, LARGE SAMPLE RESULTS*

by

Will Gersch
University of Hawaii

Technical Report No. 30

January 1977

PREPARED UNDER CONTRACT N00014-75-C-0442

(NR-042-034)

OFFICE OF NAVAL RESEARCH

Theodore W. Anderson, Project Director

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

*Supported in part by NSF Grant ENG 74-09883 at the University of Hawaii.

DISCRIMINATION BETWEEN STATIONARY GAUSSIAN PROCESSES, LARGE SAMPLE RESULTS*

by

Will Gersch
University of Hawaii

1. Introduction.

The problem considered is that of discriminating between stationary Gaussian random processes. That is, assume that a sample function $y = (y(1), \dots, y(n))$ is hypothesized to be a sample function from one of two alternative processes, with each process characterized as a zero mean covariance matrix function stationary Gaussian process. The objective is to determine the structure of the minimum classification error decision procedure and to compute the probability of misclassification. This work was primarily stimulated by a paper by Grenander [1] in which it was demonstrated that for scalar processes the probability of classification error decreased geometrically with n the number of observations. That result was achieved by Laplace's method for the evaluation of integrals in terms of the limiting distribution of the eigenvalues of a block Toeplitz matrix. The results in this paper are achieved by elementary moment analysis methods.

*Supported in part by NSF Grant ENG 74-09883 at the University of Hawaii.

2. Analysis.

1. Structure of the decision procedure: Assume that the observation matrix $y = (y(1), \dots, y(n))$ of the d -component vectors $y(j)$, $j = 1, \dots, n$ is a sample function from one of two alternative finitely generated zero mean stationary completely non-deterministic random processes. Introduce the notation, H_m , $m = 1, 2$, to denote the alternative hypotheses and assume that the a priori probability of H_1 and H_2 are equal. Then, the minimum probability of error decision procedure to distinguish the alternative hypotheses is to compute $2 \times \log$ -likelihood ratio (LR) statistic $L_n(y)$,

$$L_n(y) = 2 \ln \frac{f^1(y)}{f^2(y)} = 2 \ln \frac{f^1(y(1) \dots y(n))}{f^2(y(1) \dots y(n))} \quad (1)$$

and to decide H_1 : if $L_n(y) \geq 0$ and H_2 otherwise. In Equation (1) the superscripted quantities $f^m(y)$ denote the probability density function of the observation vector y under the m -th hypothesis.

In general the n vector of dependent normal observations $y = (y(1), \dots, y(n))$ can be orthogonalized to the form

$$\begin{aligned} f(y(1), \dots, y(n)) &= f(e(1), \dots, e(n)) = \\ &= (2\pi)^{-\frac{nd}{2}} \prod_{t=1}^n \left| \sum_t \right|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (e'(t) \sum_t^{-1} e(t)) \right]. \end{aligned} \quad (2)$$

In Equation (2) and the following unless otherwise identified, lower case letters denote vectors and upper case letters denote matrices,

an apostrophe indicates matrix transposition and $|A|$ denotes the determinant of matrix A . A Gram-Schmidt procedure is one satisfactory way to accomplish the orthogonalization. Inverse filtering and least squares prediction methods are alternative interpretations of the orthogonalizations [2]. At this point it is helpful to interpret Equation (2) in the light of some more particular structure and notation.

Using the Wold representation theorem, the assumption that $y(t)$ is completely non-deterministic is equivalent to the statement that it can be represented as

$$y(t) = \sum_{i=0}^{\infty} h_i \epsilon(t-i), \quad E[\epsilon(t)] = 0, \quad E[\epsilon(t)\epsilon'(s)] = V\delta_{t,s}. \quad (3)$$

That is, the $\{y(t)\}$ process can be thought of as the output of a linear filter with matrix impulse response $\{h_t\}$ whose input is the stationary zero mean uncorrelated sequence $\{\epsilon(t)\}$. Corresponding to the $y(t)$ process, a zero mean "residual" process $\{e(t)\}$ can be identified as the process that is produced by a filter inverse to $\{h_t\}$ acting on the $y(t)$ process. Equivalently $\{e(t)\}$ can be interpreted as the process representing the difference between $y(t)$ and its least squares estimate $\hat{y}(t)$. Let the sequence $e(1), \dots, e(n)$ in Equation (2) be a finite sample function of the $e(t)$ process. In general as identified in Equation (2) the sequence $e(1), \dots, e(n)$ need not have stationary covariances.

Identify the Wold, moving average (MA) or innovations representation of Equation (3) for the different classes of processes considered

in the operator notation

$$y^m(t) = (h^m) \epsilon^m(t), \quad E[\epsilon^m(t) \epsilon^m(s)^*] = V_m \delta_{t,s} \quad m = 1, 2. \quad (4)$$

Then from the inverse filter point of view, in operator notation the inverse of the k -th process, $k = 1, 2$, may be written

$$e^k(t) = (h^k)^{-1} y(t) = (A^k) y(t) \quad k = 1, 2. \quad (5)$$

(The assumption that $y(t)$ is finitely generated implies that the filter can be thought of as computed from a realization of the $y(t)$ unknown parameters of the $y(t)$ process.) Now consider the least squares estimator interpretation of the orthogonalization of the y sequence. It is known that $\hat{y}(t)$ the least squares estimator of $y(t)$ is the conditional mean of $y(t)$ given the past of $y(t)$ and that for Gaussian processes the estimator is a linear function of the data. Thus

$$\hat{y}(t) = E\{y(t) | y(t-1), \dots\} = \sum_{i=1}^{\infty} A_i y(t-i) \quad (6)$$

where $\{A_i\}$ are $d \times d$ coefficient matrices. Thus identify the residual process $e(t)$ by

$$e(t) = y(t) - \hat{y}(t) = \sum_{i=0}^{\infty} A_i y(t-i), \quad A_0 = I. \quad (7)$$

The infinite autoregressive (AR) representation of the residual process in Equations (7) and (5) as the inverse filter is purely formal and extremely convenient. It does not exclude alternative Markovian or Kalman filter type interpretations. When the $y(t)$ data sequence

is finite, the AR coefficient matrices become functions of time,

A_{it} .

As Grenander, our concern is with an asymptotic theory. Thus as n increases and the effects of initial conditions diminish, the residual process tends toward a stationary process. That is, the time indexed coefficient matrices A_{it} and Σ_t tend to the constant matrices A_1 and V respectively. Applying the forgoing considerations to the distribution of $f^1(y)$, $f^2(y)$ in the likelihood ratio statistic $L_n(y)$ we obtain:

Lemma 1. The asymptotic log-likelihood ratio statistic for distinguishing between alternative stationary Gaussian time series is

$$L_n(y) = 2\ell n \frac{f^1(y)}{f^2(y)} = n \log \frac{|V_2|}{|V_1|} + \sum_{t=1}^n e^2(t)' V_2^{-1} e^2(t) - \sum_{t=1}^n e^1(t)' V_1^{-1} e^1(t) \quad (8)$$

In Equation (8) $e^m(t)$ is the residual at time t obtained by filtering the observed $y(t)$ process with the m -th model inverse filter, $m = 1, 2$. The optimal decision procedure thus involves the application of an inverse of filter to the observed data sequence $y(1), \dots, y(n)$. Calculation of the LR statistic in Equation (8) and application of the decision rule, decide H_1 if $L_n(y) \geq 0$ and decide H_2 otherwise. The optimum decision procedure structure is illustrated in Figure 1.

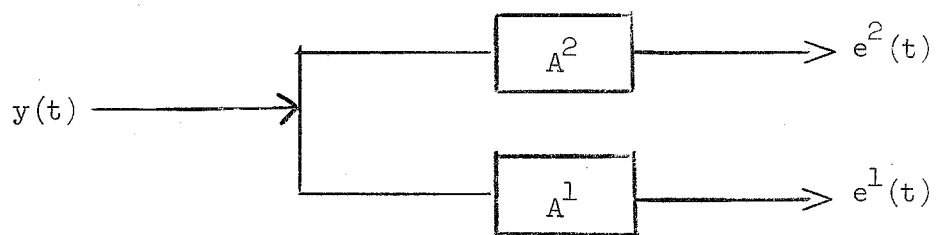


Figure 1. The structure of the LR decision procedure.

2. Statistics of the Log-Likelihood ratio test:

Lemma 2. The log-likelihood ratio test statistic is asymptotically normally distributed with conditional means and variances μ_m , σ_m^2 $m = 1, 2$;

$$L_n(y)|H_m \sim N(\mu_m, \sigma_m^2), \quad m = 1, 2. \quad (9)$$

(Explicit formulas for μ_m , σ_m^2 are in Lemmas 3 and 4).

Proof. The LR statistic, Equation (8) is the sum of sums of dependent quadratic terms. Under conditions on the rate of decay of dependence, that are difficult to ascertain, each of the sum of terms tends toward a normal distribution. On the other hand, each of the sum of terms is in the form

$$\begin{aligned} \sum_{t=1}^n e(t)' V^{-1} e(t) &= \sum_{t=1}^n \text{tr}(e(t)e(t)' V^{-1}) = \text{tr}\left(\sum_{t=1}^n e(t)e(t)'\right) V^{-1} \\ &= n \text{tr}[Cee(0) V^{-1}] \end{aligned} \quad (10)$$

In Equation (10) $\text{tr}[A]$ is the trace of matrix A , $\text{tr}[AB] = \text{tr}[BA]$, $Cee(0)$ is the sample covariance matrix of a residual process calculated at zero lag and V is a theoretical (constant) residual variance matrix. Based upon the earlier work of Diananda [3] and Walker [4] and following Hannan [5] and Anderson [6] the elements of $Cee(0)$ are asymptotically normally distributed. Thus $Cee(0)$ and $Cee V^{-1}$ have multivariate normal distributions and asymptotically $\text{tr}(Cee(0) V^{-1})$ has a limiting scalar distribution. Following Hannan, a sufficient condition for

normality of the sample covariance is that if $\{h_t\}$ is the impulse response matrix of the $e(t)$ process that $\sum_{t=0}^{\infty} |th_t| < \infty$. Explicit formula for $\{h_t\}$ are given following Lemma 3 and this condition is seen to be trivially satisfied. Lemma 2 expresses the situation of preimary interest, the distribution of the LR statistic under the alternative hypotheses. The conditional means and variances of the limiting distribution of the log-likelihood ratio statistic are computed next.

Lemma 3. Time and frequency domain formula for the conditional mean and variance of the LR statistic are:

$$E[L_n(y)|H_1] = 2 \int f^1(y) \ln \frac{f^1(y)}{f^2(y)} dy = \mu_1 \triangleq 2n I(1:2) \quad (11)$$

$$= n \left[\ln \frac{|V_2|}{|V_1|} + \text{tr} \int_{-1/2}^{1/2} S_1(f) S_2(f)^{-1} df - d \right] \quad (a)$$

$$= n \left[\ln \frac{|V_2|}{|V_1|} + \text{tr} \left[\sum_{j=0}^{\infty} (h_j^{1,2} V_1 h_j^{1,2'}) V_2^{-1} \right] - d \right] \quad (b)$$

$$E[L_n(y)|H_2] = 2 \int f^2(y) \ln \frac{f^1(y)}{f^2(y)} dy = \mu_2 \triangleq -2n I(2:1)$$

$$= n \left[\ln \frac{|V_2|}{|V_1|} - \text{tr} \int_{-1/2}^{1/2} S_2(f) S_1(f)^{-1} df + d \right] \quad (c)$$

$$= n \left[\ln \frac{|V_2|}{|V_1|} - \text{tr} \left[\sum_{j=0}^{\infty} (h_j^{2,1} V_2 h_j^{2,1'}) V_1^{-1} \right] + d \right] \quad (d)$$

In Equation (11) by definition, $I(1:2)$ is the Kullback-Liebler [7] measure of the amount of information per observation to decide H_1

instead of H_2 when H_1 is true, $S_m(f)$ $m = 1, 2$, are the power spectral densities of the $\{y^1(t)\}$ and $\{y^2(t)\}$ process respectively and $h_t^{k,m}$ is the impulse response of the residual $e^{k,m}(t)$ process, ($h^{k,m} = A^m h^k$) the process formed by acting on the process $y^k(t)$ with the model corresponding to the inverse of the m -th process. The time and frequency domain formulas in Equation (11) are new.

Proof. Consider the conditional expectations of each of the terms in Equation (8) separately. Then,

$$E\left[\sum_{t=1}^n e^{1'}(t) V_1^{-1} e^1(t) | H_1\right] = \text{tr}\left[\left(\sum_{t=1}^n e^1(t) e^{1'}(t) | H_1\right) V_1^{-1}\right] = nd, \quad (12)$$

because

$$E\left(\sum_{t=1}^n e^1(t) e^{1'}(t) | H_1\right) = n E[Ce^1 e^{1'}(0)] = nV_1.$$

Then, too,

$$E\left[\sum_{t=1}^n e^{2'}(t) V_2^{-1} e^2(t) | H_1\right] = \text{tr} E\left[\left(\sum_{t=1}^n e^2(t) e^{2'}(t) | H_1\right) V_2^{-1}\right]. \quad (13)$$

Now by definition of the $e^2(t)$ process

$$e^2(t) | H_1 = (A^2)y^1(t) = (A^2h^1)\epsilon^1(t) = (h^{1,2})\epsilon^1(t) \quad (14)$$

$$= \sum_{i=0}^{\infty} h_i^{1,2} \epsilon^1(t-i).$$

Direct expansion of the operator expressions in Equation (14) yields the result

$$h_t^{k,m} = [h_t^k + \sum_{i=1}^{\infty} A_i^m h_{t-i}^k] , \quad t = 0, 1, \dots \quad (15)$$

Digression: It is convenient here to think of A_i^m as the AR process coefficient matrices in the representation of $y^m(t)$. Thus we observe that $h_0^{k,m} = I$, $h_t^{k,k} = I\delta_t$ and $\text{Var}(e^{k,m}(t)) \geq \text{Var}(e^{k,k}(t))$ with equality if and only if $m = 1$. These conclusions are a direct consequence of the relationship between the AR and MA representations of a stationary process. Thus in the AR and MA representations

$$y^k(t) = (h^k)\epsilon^k(t) , \quad (A^k)y^k(t) = \epsilon^k(t) . \quad (16)$$

From this we obtain the useful recursive relationships for A_i^k, h_i^k

$$\begin{aligned} (A^k)(h^k) &= I , \quad h_t^k = - \sum_{i=1}^{\infty} A_i^k h_{t-i}^k , \quad t = 1, 2, \dots \\ (h^k)(A^k) &= I , \quad A_t^k = - \sum_{i=1}^{\infty} h_t^k A_{t-i}^k , \quad h_0^k = A_0^k = I . \end{aligned} \quad (17)$$

Returning to the proof of the lemma,

$$\begin{aligned}
E\left(\sum_{t=1}^n e^{2'}(t) V_2^{-1} e^2(t) | H_1\right) &= \text{tr} \sum_{t=1}^n E[A^2 h^1 \epsilon^1(t) \epsilon^{1'}(t) h^{1'} A^{2'} V_2^{-1}] \\
&= n \text{tr}[(h^1 V_1 h^{1'}) (A^{2'} V_2^{-1} A^2)] \\
&= n \text{tr}\left[\int_{-1/2}^{1/2} S_1(f) S_2(f)^{-1} df\right]
\end{aligned} \tag{18}$$

In Equation (18) we made repeated use of the matrix interchange property of the trace operation, and the fact that for stationary process the covariance function and spectral densities are Fourier transform pairs. We also used the formulas for the spectral density of a process $S(f) = h V h^\dagger = A^{-1} V A^{-1\dagger}$ where h and A are the usual polynomial matrix frequency domain operators, [2] and the dagger \dagger denotes complex conjugate transpose. The combination of the methods and results in Equations (13)-(18) yield the frequency domain formulas for the conditional mean LR statistic in Equation (11a and 11c). Shumway and Unger [8] proved a result similar to Equation (16) for the scalar case via a limiting distribution of the eigenvalues of a Toeplitz matrix development.

The frequency domain formulas for the Kullback-Liebler numbers, Equations (11a and 11c) may be computed by a variety of techniques. The corresponding time domain formulas, Equations (11b and 11d) are very satisfactorily computed by an approximating sum. The time domain formulas may be obtained directly starting with Equation (14). Using the matrix interchange property of the trace, the formula in Equation (13) for $e^2(t) | H_1$, and the fact that $E[e^1(t) e^{1'}(t) | H_1] = V_1$,

$$\begin{aligned}
E\left[\sum_{t=1}^n e^{2'}(t)V_2^{-1}e^2(t)|H_1\right] &= \sum_{t=1}^n \text{tr} E(e^{1,2}(t)V_1e^{1,2'}(t))V_2^{-1} \\
&= n \text{tr}\left[\sum_{j=0}^{\infty} (h_j^{1,2}V_1h_j^{1,2'})V_2^{-1}\right].
\end{aligned} \tag{19}$$

A side point is that a useful interpretation of the Kullback-Liebler measure, taken directly from the definition in Equation (11) is,

$$f^2(y(1), \dots, y(n)|H_1) = \exp -nI(1:2)f^1(y(1), \dots, y(n)|H_1) \tag{20}$$

That is, the value of the likelihood (incremental probability) of the observed sequence $y(1), \dots, y(n)$ under the assumption of model 2 when process 1 is true, is on the average $\exp -nI(1:2)$ times the value of the likelihood of that sequence under the assumption of model 1, the true model of the observed process.

LEMMA 4. The conditional variances of the LR statistics are

$$\sigma_1^2 = 2n[d^2 + \sum_{r=0}^{n-1} (1 - \frac{r}{n})\{\text{tr}^2(\Gamma_r^{1,2}) + \text{tr}^2(\Gamma_{-r}^{1,2}) - \text{dtr}(h_r^{1,2})\text{tr}(h_r^{1,2'}V_2^{-1}V_1)\}] \tag{21}$$

$$\sigma_2^2 = 2n[d^2 + \sum_{r=0}^{n-1} (1 - \frac{r}{n})\{\text{tr}^2(\Gamma_r^{2,1}) + \text{tr}^2(\Gamma_{-r}^{2,1}) - \text{dtr}(h_r^{2,1})\text{tr}(h_r^{2,1'}V_1^{-1}V_2)\}]$$

where

$$\Gamma_r^{k,m} = \left[\sum_{j=0}^{\infty} h_j^{k,m} V_k h_{j+r}^{k,m'} \right] V_m^{-1}, \quad \Gamma_{-r}^{k,m} = (\Gamma_k^{k,m})' \tag{22}$$

Proof: This lemma is a computational result that requires computation of the conditional variance of the summed terms in the LR statistic in Equation (8) and the covariance of those terms. These are only second moment computations, they follow the spirit and the technique of the computations for the previous lemma. Therefore the obvious algebraic details are omitted. Consider computation of the conditional variance of the first summed term in Equation (8). Using the formula $\text{Var}(X) = E[X^2] - E^2[X]$ we write

$$E\left[\sum_{t=1}^n e^{2'}(t) V_2^{-1} e^2(t) | H_1\right] = E\left[\sum_{r=1}^d \sum_{j=1}^d \sum_{t=1}^n \sum_{s=1}^n \tilde{e}_j(t) \tilde{e}_j(t) \tilde{e}_r(s) \tilde{e}_r(s)\right]. \quad (23)$$

To obtain Equation (23) we employ the upper-lower triangular factorization of $V_2 = L(V_2)U(V_2)$, $V_2^{-1} = U^{-1}(V_2)L^{-1}(V_2)$ and identify the components $\tilde{e}_j(t)$ via the multiplication $e^{2'}(t)U^{-1}(V_2) = [\tilde{e}_1(t), \dots, \tilde{e}_d(t)]$. Then, recall the known result for normally distributed random variables

$$E[X_1 X_2 X_3 X_4] = E[X_1 X_2]E[X_3 X_4] + E[X_1 X_3]E[X_2 X_4] + E[X_1 X_4]E[X_2 X_3]. \quad (24)$$

Apply this result in Equation (23), recombine terms and employ direct evaluation to obtain

$$\begin{aligned} \text{Var}\left[\sum_{t=1}^n e^{2'}(t) V_2^{-1} e^2(t) | H\right] &= 2 \sum_{t=1}^n \sum_{s=1}^n (\text{tr } E[e^2(s) e^2(t)' V_2^{-1}])^2 \\ &= 2n \sum_{r=1}^d \left(1 - \frac{r}{n}\right) \{\text{tr}^2(\Gamma_r^{1,2}) + \text{tr}^2(\Gamma_{-r}^{1,2})\} \end{aligned} \quad (25)$$

where $\Gamma_r^{k,m}$, $\Gamma_{-r}^{k,m}$ are as defined in Equation (23). Similarly,

$$\text{Var}\left(\sum_{t=1}^n e^1(t)' V_1^{-1} e^1(t) | H_1\right) = 2 \sum_{t=1}^n \sum_{s=1}^n (\text{tr } E[e^1(s) e^1(t)' V_1^{-1}])^2 = 2nd^2. \quad (26)$$

The covariance terms may be evaluated as

$$\begin{aligned} \text{Cov}\left(\sum_{t=1}^n e^2(t)' V_2^{-1} e^2(t) \sum_{s=1}^n e^1(s)' V_1^{-1} e^1(s) | H_1\right) \\ = 2 \sum_{t=1}^n \sum_{s=1}^n E[e^1(s) e^1(t)' h_r^{1,2'} V_2^{-1}] E[h_r^{1,2} e^1(t) e^1(s)' V_1^{-1}] \quad (27) \\ = 2nd \sum_{r=0}^{n-1} \left(1 - \frac{r}{n}\right) \text{tr}(h_r^{1,2}) \text{tr}(h_r^{1,2'} V_2^{-1} V_1). \end{aligned}$$

THEOREM. The asymptotic probability of classification error between stationary Gaussian processes is bounded exponentially with n , the number of observations.

Proof. By virtue of the fact that the LR statistic is asymptotically normal, consider direct evaluation of classification error under H_1 .

Assume $\mu_1 > \mu_2$. Then

$$P_n(\text{error} | H_1) = \int_{-\infty}^0 L_n(y) | H_1 dy = \Phi\left(-\frac{\mu_1}{\sigma_1}\right). \quad (28)$$

In Equation (28), by elementary considerations $\mu_1 > 0$. Since μ_1 and σ_1^2 are linearly proportional to n , $\alpha = \frac{\mu_1}{\sigma_1} = \sqrt{n} c_1 > 0$. Thus

$$\Phi\left(\frac{-\mu_1}{\sigma_1}\right) = 1 - \Phi(\alpha) \leq \frac{1}{\sqrt{2\pi} \alpha} \exp\left(-\frac{\alpha^2}{2}\right) = \frac{1}{\sqrt{2\pi} nc_1} \exp\left(-\frac{nc_1^2}{2}\right). \quad (29)$$

The bound is obtained by integration by parts (i.e., Mills' ratio).

Similar results are obtained for $P_n(\text{error}|H_2)$ and the alternative assumption that $\mu_1 < \mu_2$.

Discussion.

This paper presents a general result on the asymptotic theory of the discrimination between multivariate stationary Gaussian random processes. The result that the asymptotic probability of error bounded exponentially with n is stronger than that achieved by Grenander. Also important, this analysis readily suggests implementations to achieve minimum classification error between stationary random processes.

REFERENCES

- [1] Grenander, U. (1974), Large sample discrimination between two Gaussian processes with different spectra, Annals of Stat., 2, 347-352.
- [2] Whittle, P. (1963), Prediction and Regulation by Least Squares Methods, The English Universities Press.
- [3] Diananda, P.H. (1953), Some probability limit theorems with statistical applications, Proc. Camb. Phil. Soc., 49, 239-246.
- [4] Walker, A.M. (1954), The asymptotic distribution of serial correlation coefficients for autoregressive processes with dependent residuals, Proc. Camb. Phil. Soc., 50, 60-64.
- [5] Hannan, E.J. (1960), Time Series Analysis, Methuen.
- [6] Anderson, T.W. (1971), The Statistical Analysis of Time Series, Wiley.
- [7] Kullback, S. (1959), Information Theory and Statistics, Wiley.
- [8] Shumway, R.H. and Unger, A.N. (1974), Linear discriminant functions for stationary time series, Jour. Am. Stat. Assoc., 69, 948-956.

TECHNICAL REPORTS

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030 (NR-042-034)

1. "Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function," T. W. Anderson, October 1, 1969.
2. "Efficient Estimation of Regression Coefficients in Time Series," T. W. Anderson, October 1, 1970.
3. "Determining the Appropriate Sample Size for Confidence Limits for a Proportion," T. W. Anderson and H. Burstein, October 15, 1970.
4. "Some General Results on Time-Ordered Classification," D. V. Hinkley, July 30, 1971.
5. "Tests for Randomness of Directions against Equatorial and Bimodal Alternatives," T. W. Anderson and M. A. Stephens, August 30, 1971.
6. "Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order," T. W. Anderson, October 29, 1971.
7. "The Stationarity of an Estimated Autoregressive Process," T. W. Anderson, November 15, 1971.
8. "On the Inverse of Some Covariance Matrices of Toeplitz Type," Raul Pedro Mentz, July 12, 1972.
9. "An Asymptotic Expansion of the Distribution of "Studentized" Classification Statistics," T. W. Anderson, September 10, 1972.
10. "Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions," T. W. Anderson, September 28, 1972.
11. "Population Mixing Models and Clustering Algorithms," Stanley L. Sclove, February 1, 1973.
12. "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance," John James Miller, November 21, 1973.
13. "Maximum Likelihood Estimation in the Birth-and-Death Process," Niels Keiding, November 28, 1973.
14. "Random Orthogonal Set Functions and Stochastic Models for the Gravity Potential of the Earth," Steffen L. Lauritzen, December 27, 1973.
15. "Maximum Likelihood Estimation of Parameters of an Autoregressive Process with Moving Average Residuals and Other Covariance Matrices with Linear Structure," T. W. Anderson, December, 1973.
16. "Note on a Case-Study in Box-Jenkins Seasonal Forecasting of Time Series," Steffen L. Lauritzen, April, 1974.

TECHNICAL REPORTS (continued)

17. "General Exponential Models for Discrete Observations,"
Steffen L. Lauritzen, May, 1974.
18. "On the Interrelationships among Sufficiency, Total Sufficiency and
Some Related Concepts," Steffen L. Lauritzen, June, 1974.
19. "Statistical Inference for Multiply Truncated Power Series Distributions,"
T. Cacoullos, September 30, 1974.

Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)

20. "Estimation by Maximum Likelihood in Autoregressive Moving Average Models
in the Time and Frequency Domains," T. W. Anderson, June 1975.
21. "Asymptotic Properties of Some Estimators in Moving Average Models,"
Raul Pedro Mentz, September 8, 1975.
22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting,"
Mituaki Huzii, February 1976.
23. "Estimating Means when Some Observations are Classified by Linear
Discriminant Function," Chien-Pai Han, April 1976.
24. "Panels and Time Series Analysis: Markov Chains and Autoregressive
Processes," T. W. Anderson, July 1976.
25. "Repeated Measurements on Autoregressive Processes," T. W. Anderson,
September 1976.
26. "The Recurrence Classification of Risk and Storage Processes,"
J. Michael Harrison and Sidney I. Resnick, September 1976.
27. "The Generalized Variance of a Stationary Autoregressive Process,"
T. W. Anderson and Raul P. Mentz, October 1976.
28. "Estimation of the Parameters of Finite Location and Scale Mixtures,"
Javad Behboodian, October 1976.
29. "Identification of Parameters by the Distribution of a Maximum
Random Variable," T. W. Anderson and S.G. Ghurye, November 1976.
30. "Discrimination Between Stationary Gaussian Processes, Large Sample
Results," Will Gersch, January 1977.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 30	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) DISCRIMINATION BETWEEN STATIONARY GAUSSIAN PROCESSES, LARGE SAMPLE RESULTS		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Will Gersch		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0442
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, California 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-042-034)
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 436 Arlington, Virginia 22217		12. REPORT DATE January 1977
		13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; Distribution Unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Discrimination, Stationary Gaussian processes.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The structure of the asymptotic log-likelihood ratio decision procedure for the discrimination of alternative stationary zero mean multivariate Gaussian processes is developed. The log-likelihood ratio statistic is shown to be asymptotically normally distributed. New time and frequency domain formulas for the conditional mean (The Kullback-Liebler information measure) and variance of the log-likelihood ratio statistic under the alternative hypotheses are given and the probability of misclassification is shown to be bounded exponentially with n the number of observations.		

DD FORM 1473

1 JAN 73

EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)